

机器学习与量化交易实战

第七讲



看大家写的作业，读后感如图：

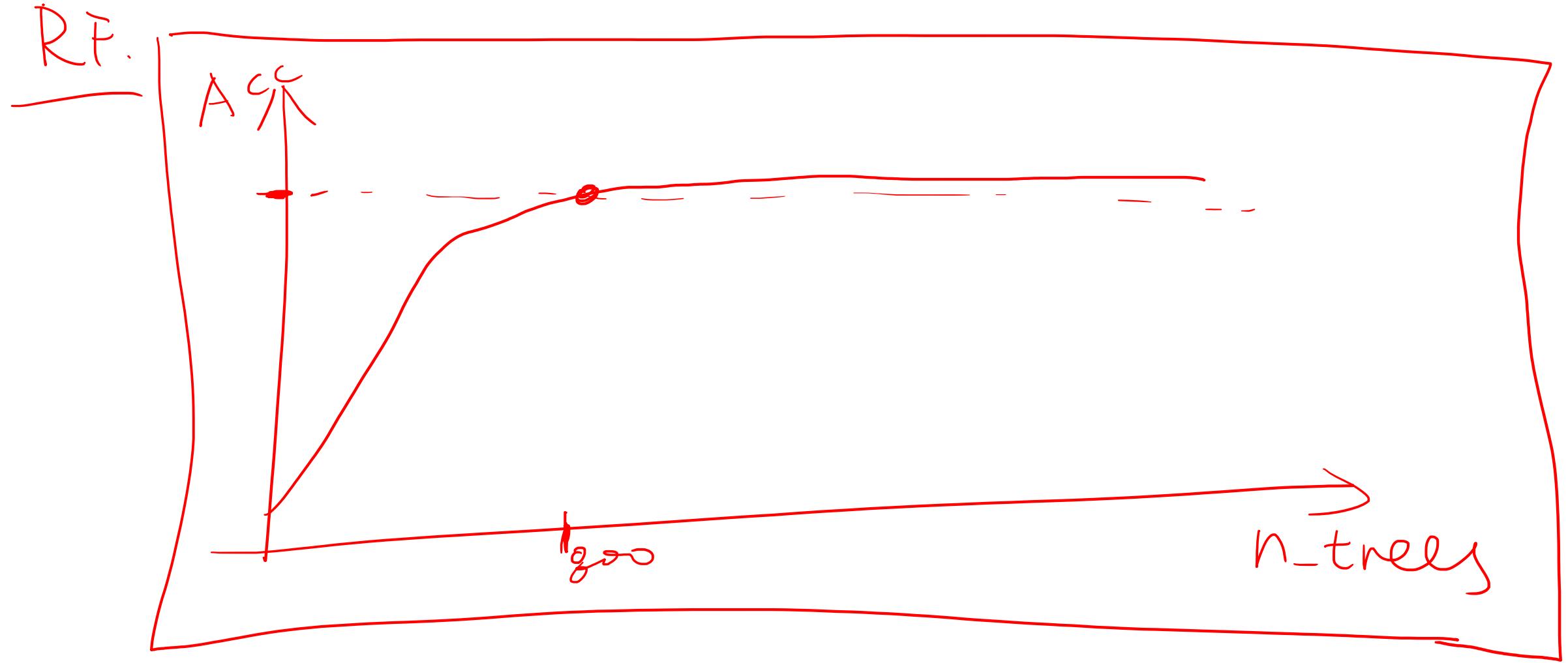


① 数据是否平衡

- e.g. Class balanced?
- Statistics
- Visualization 可视化

②

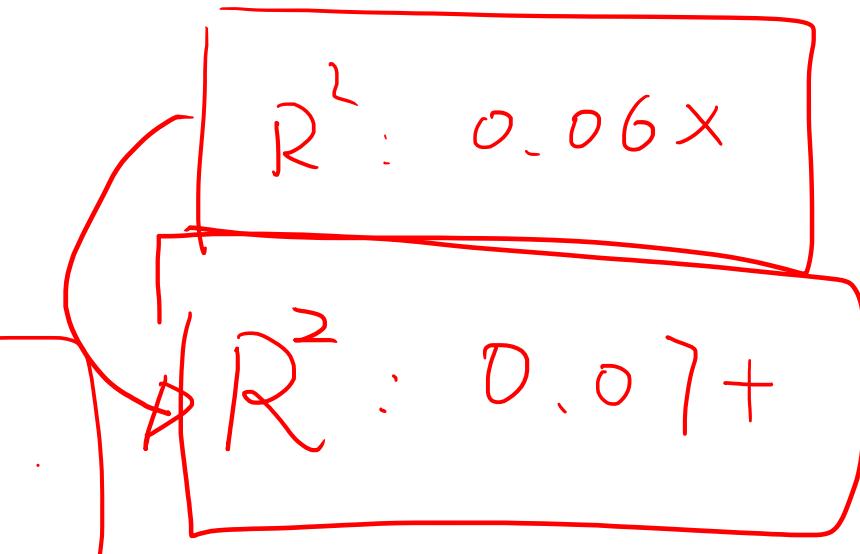
→ 若有不平衡选择 C/R



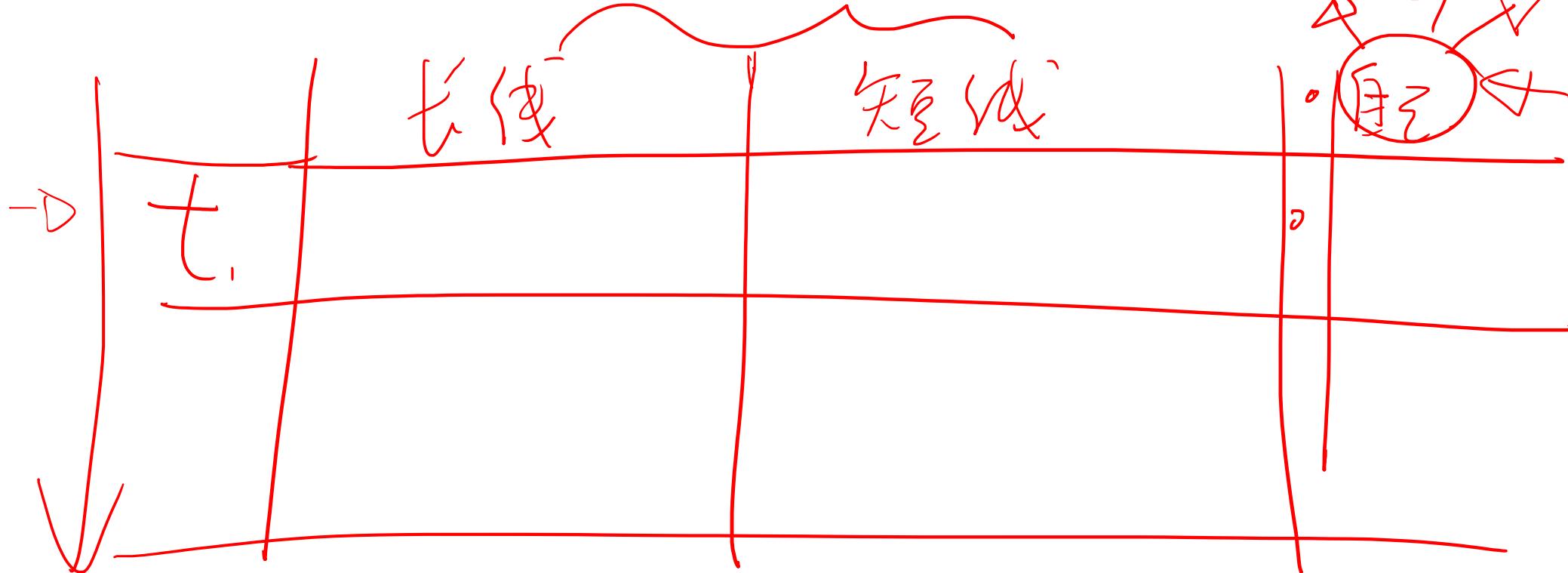
① 球面傳播.

② 多相.

③ |Prap. Space.



拉斐特 FEATURES.



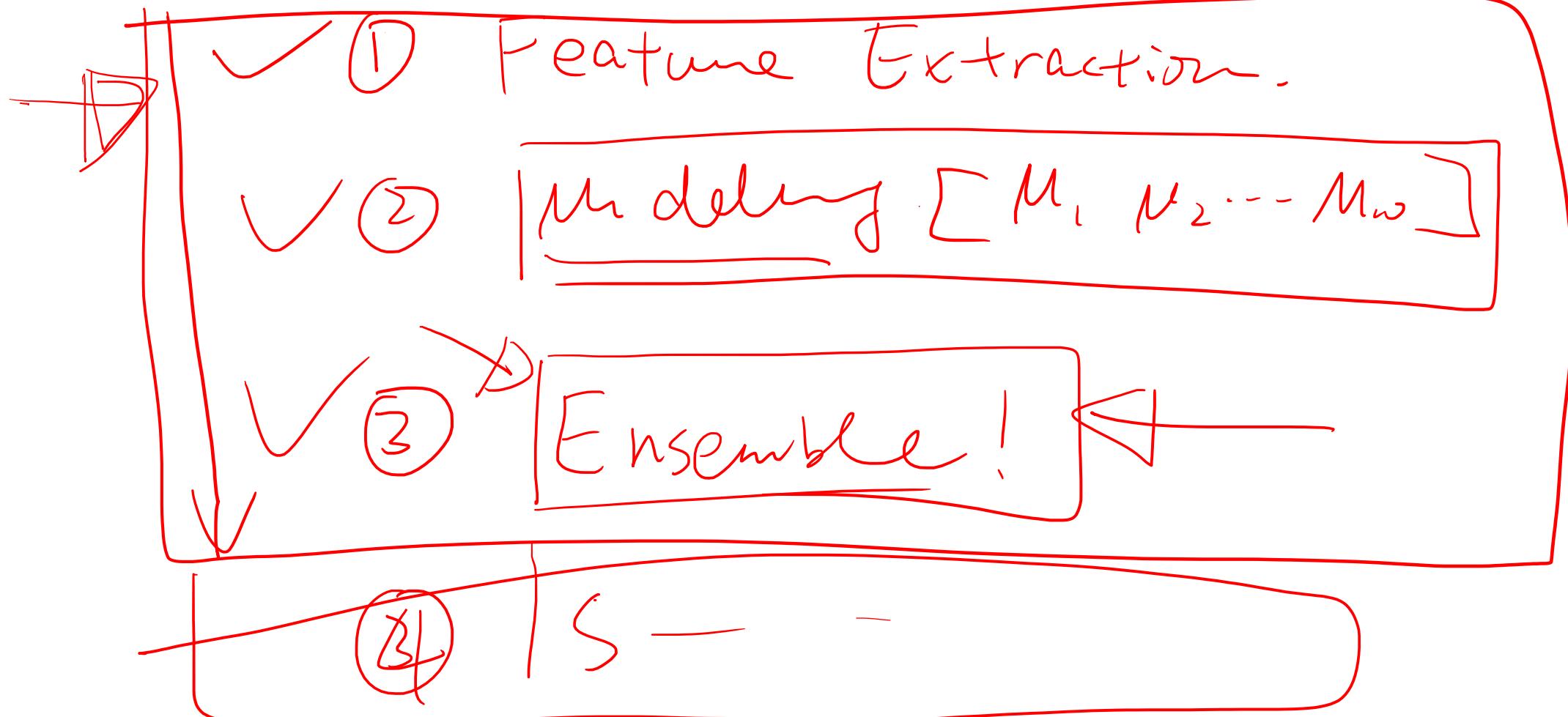
声母

ENSEMBLE

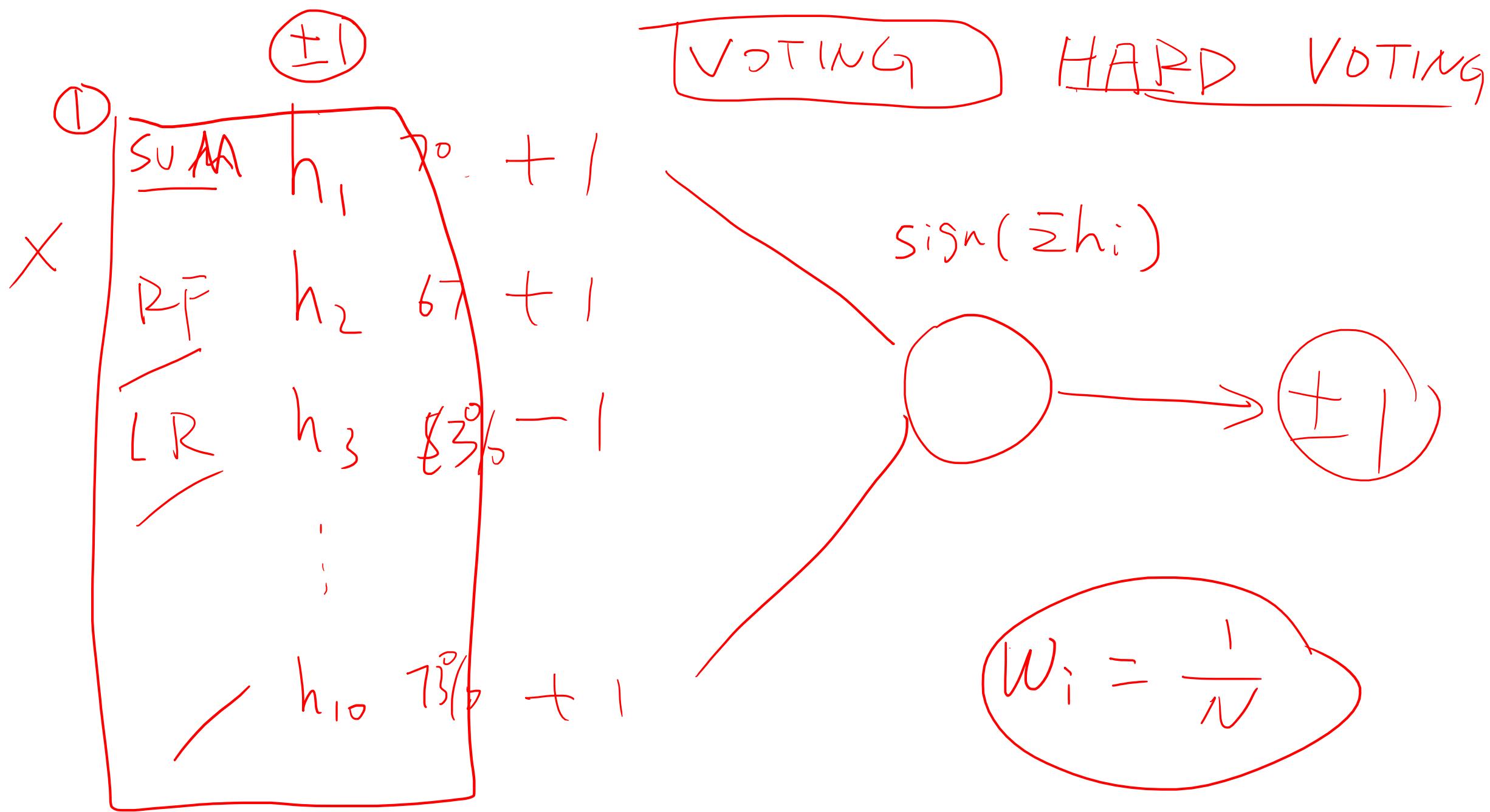
集成學習

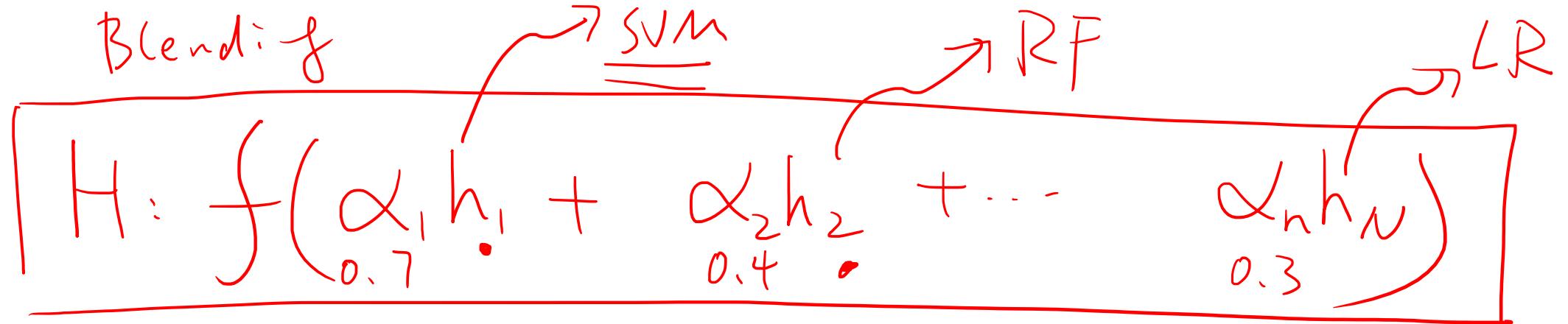
ASS - ~~SAC~~
SAVING

METHOD.



- Stacking / blending / Voting EASY
IMPORTANT
- AdaBoost ←

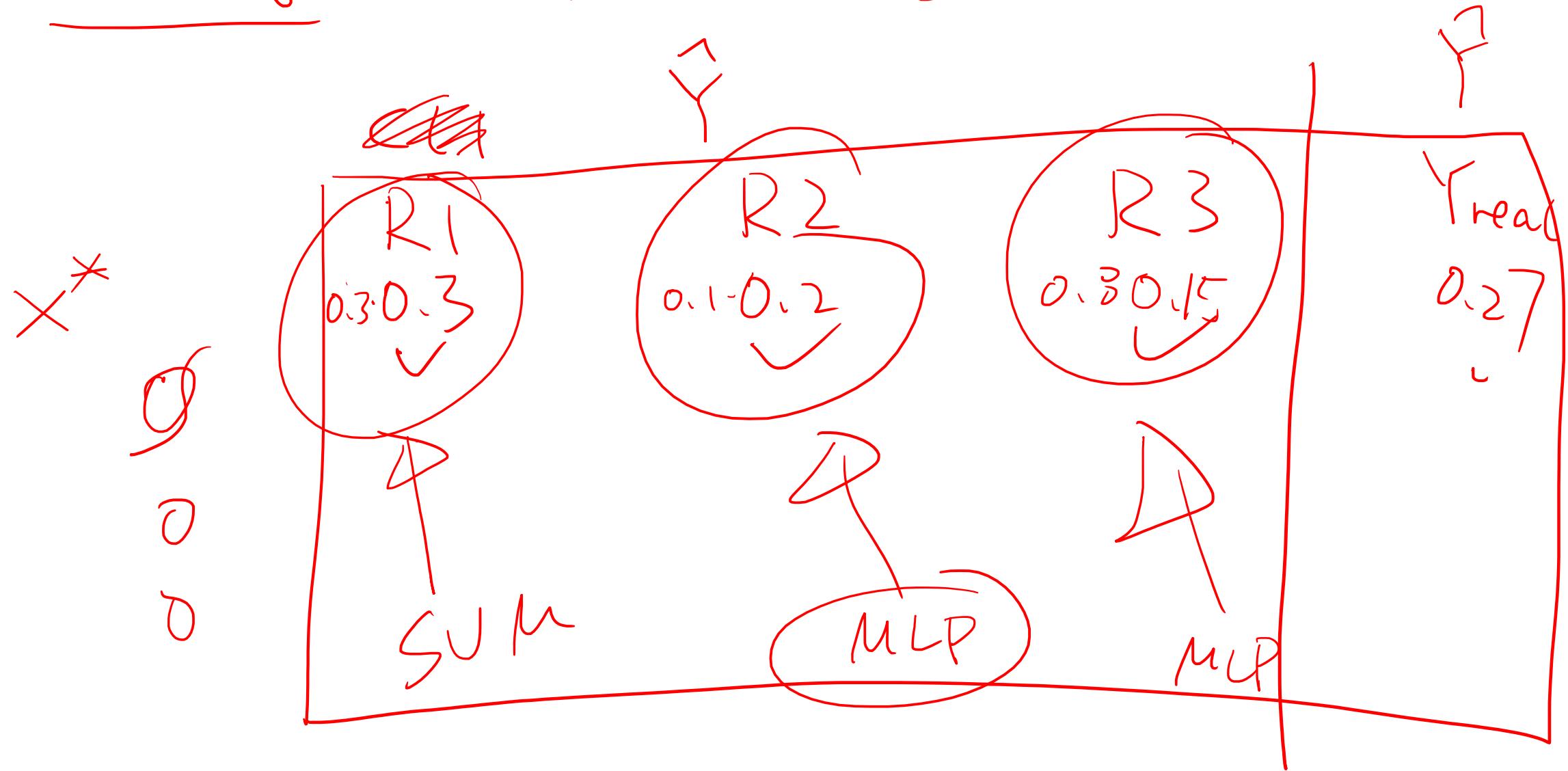


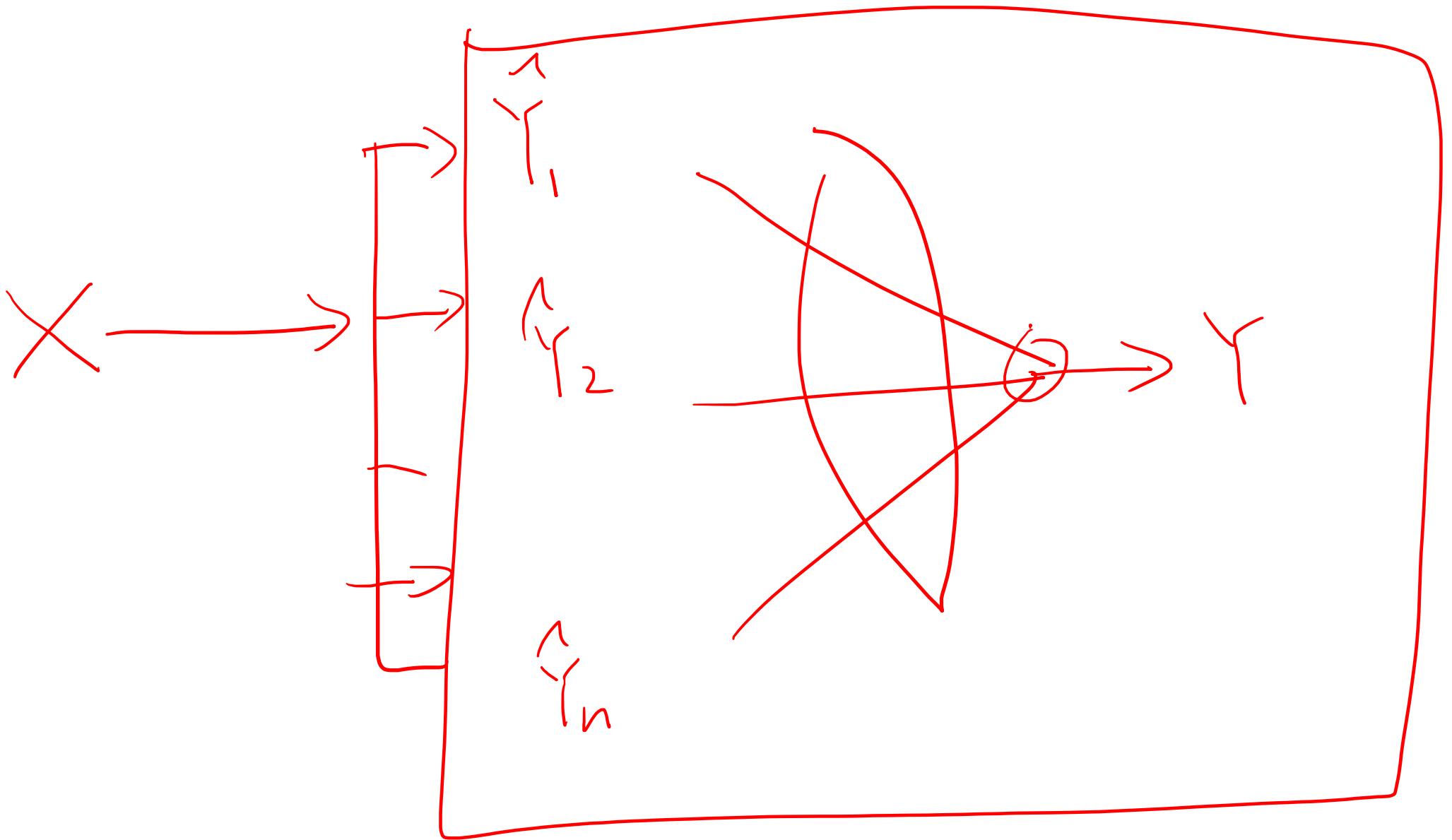


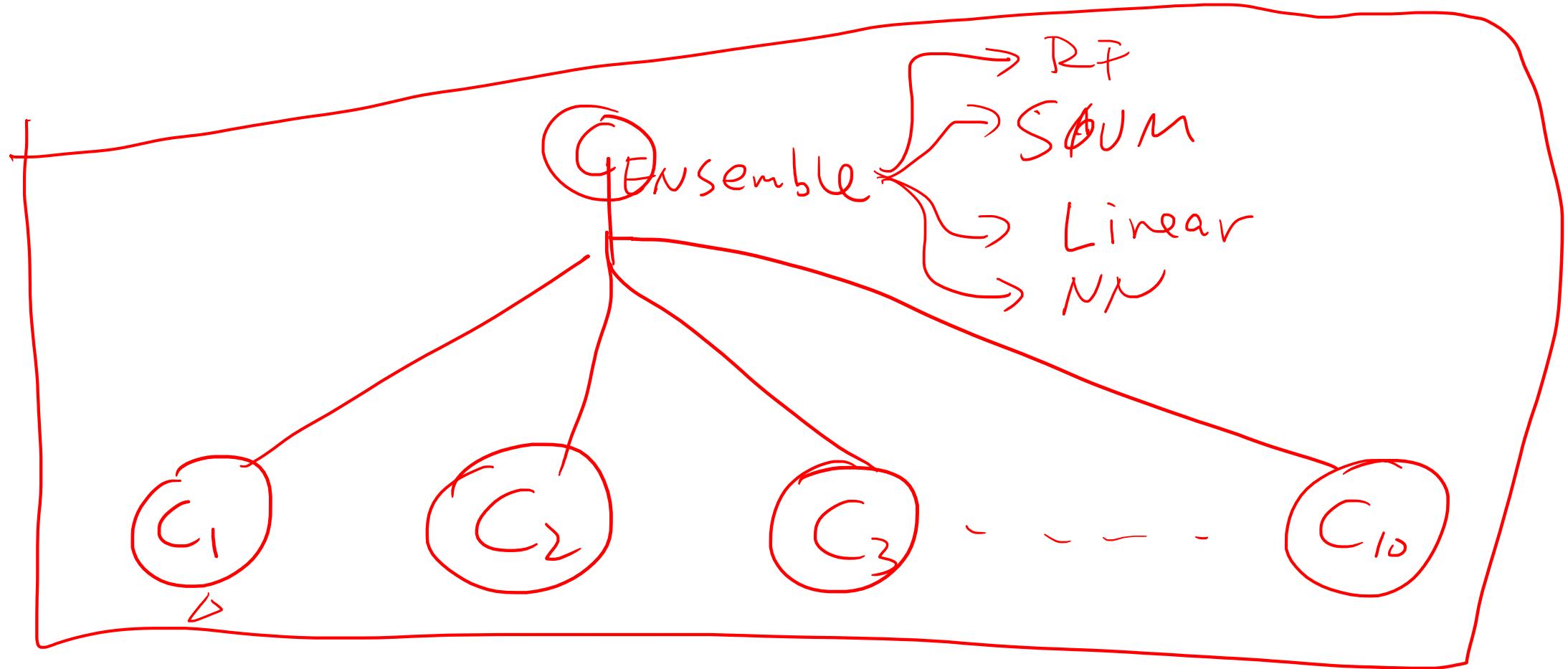
Q: How TO Define weights for clf_s

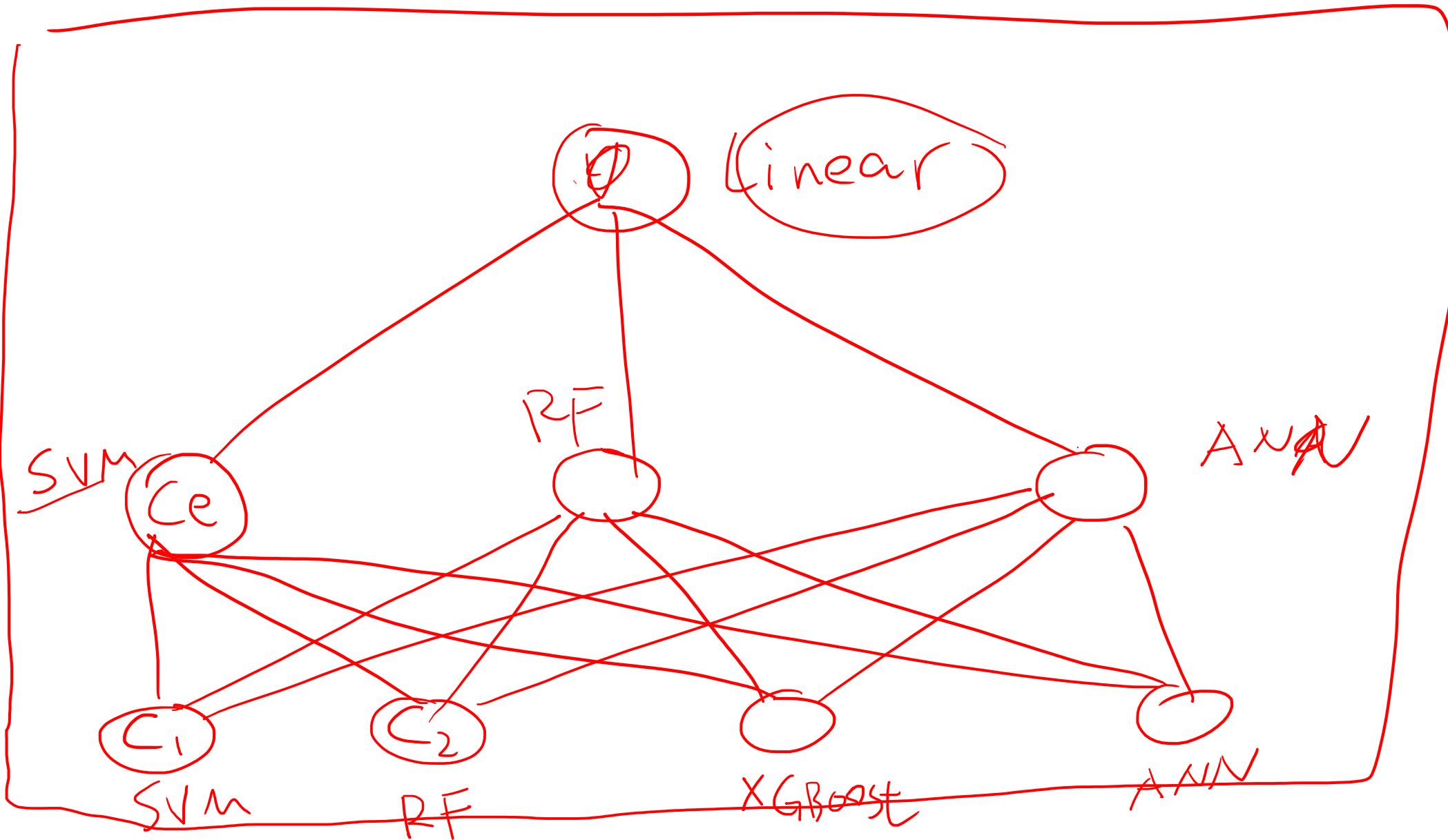
①	Acc	w
→ SUM	70%	Acc/̄Acc
LR	63%	

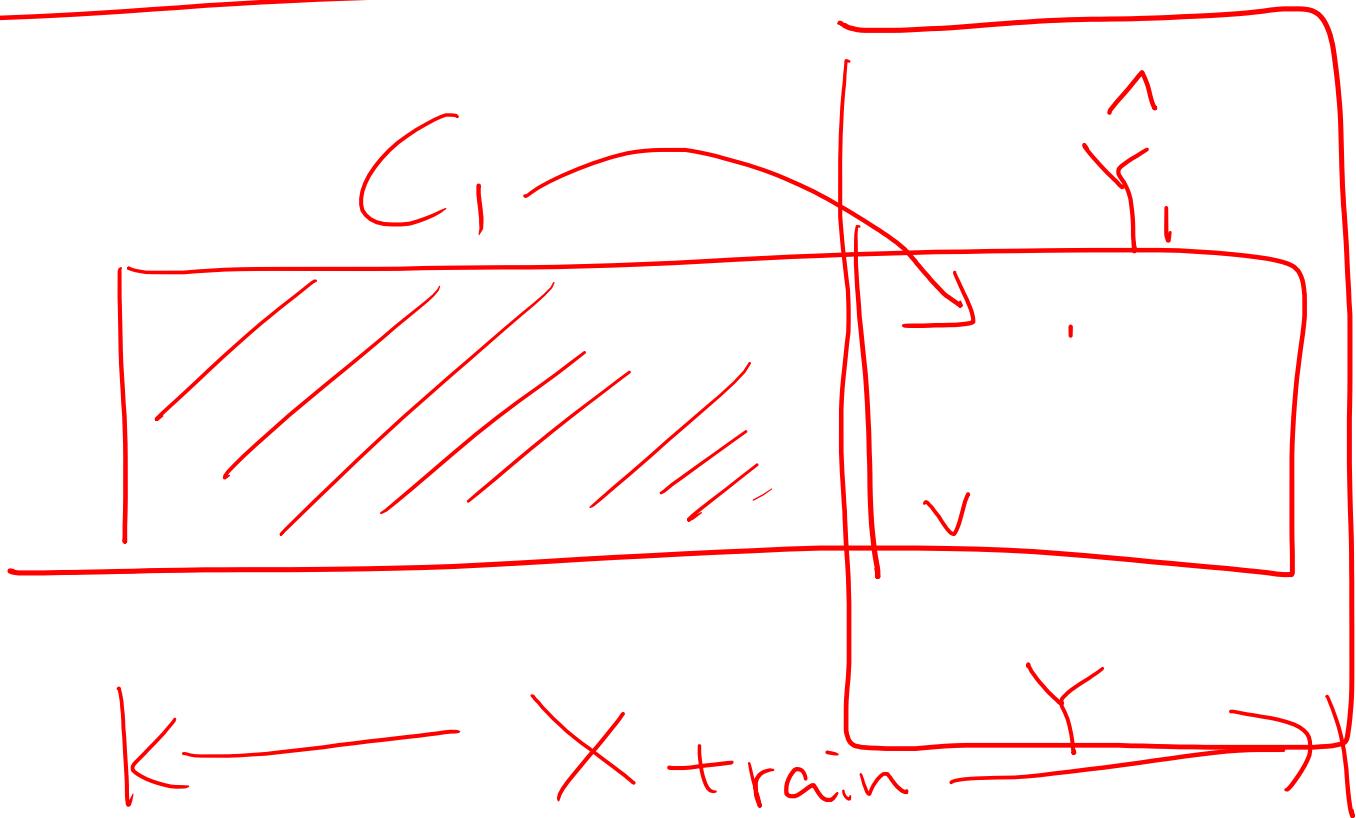
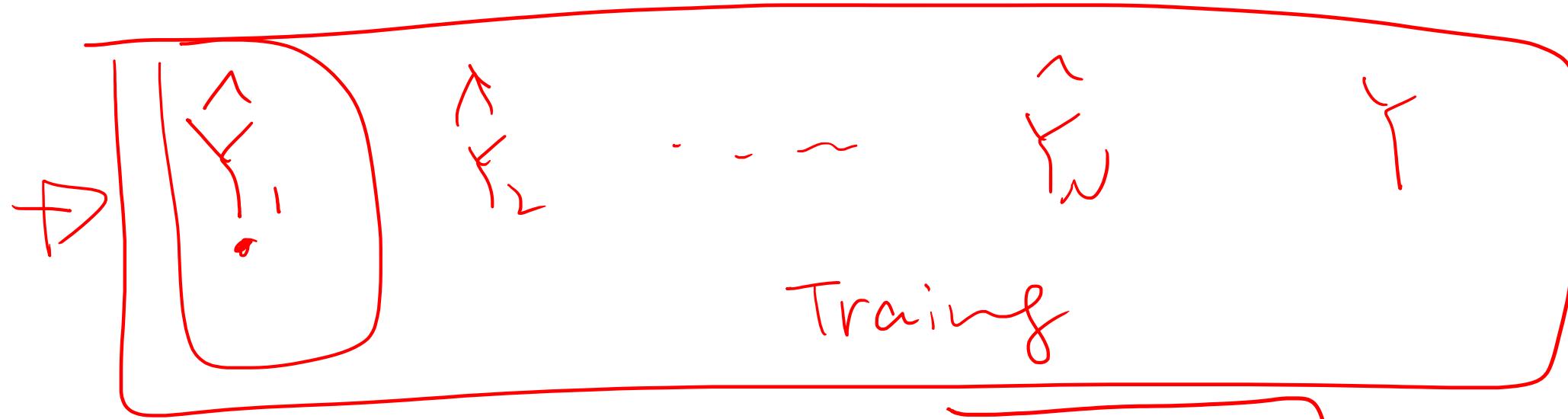
Blending: 在黑色背景的方法來做。

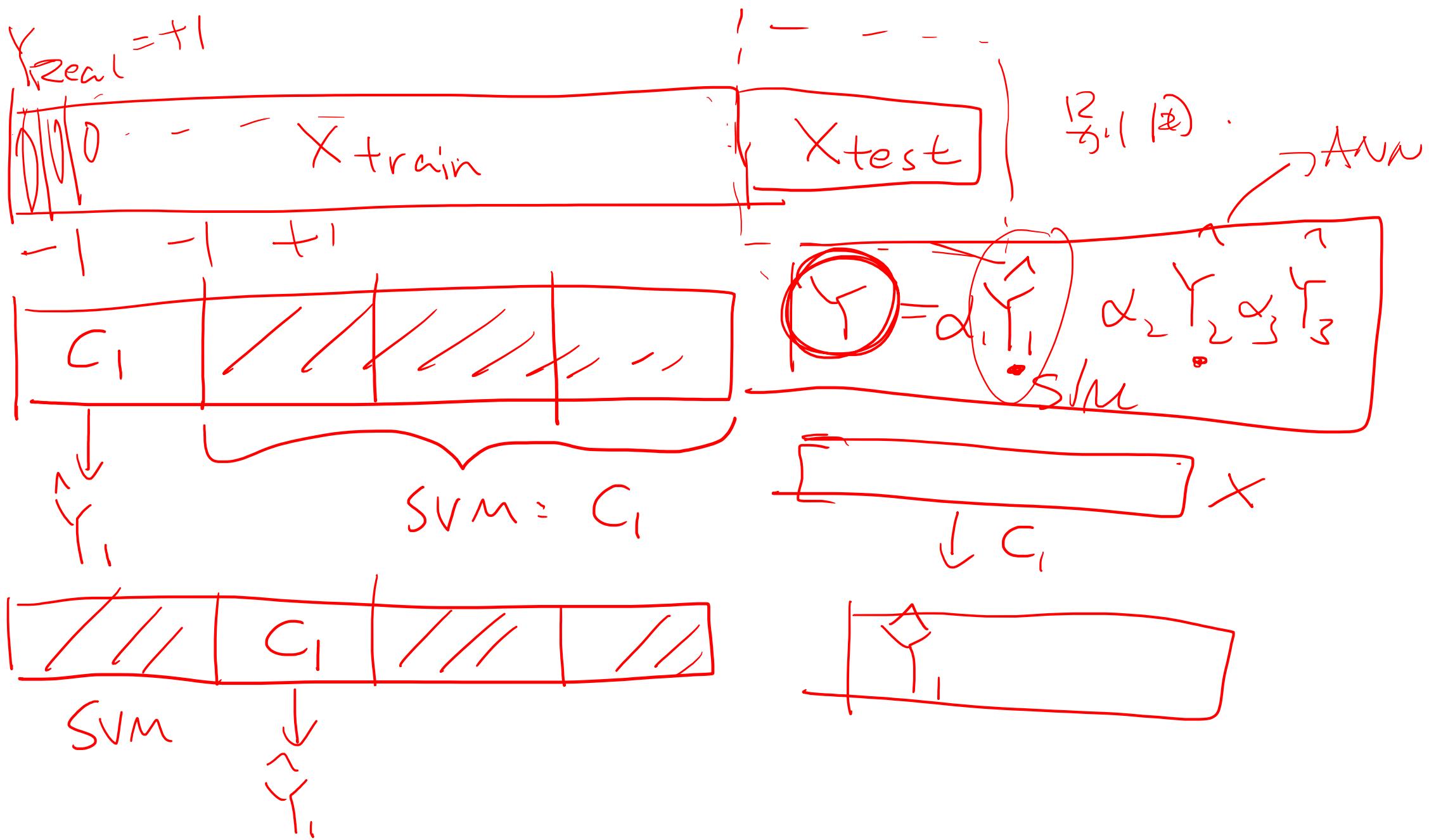


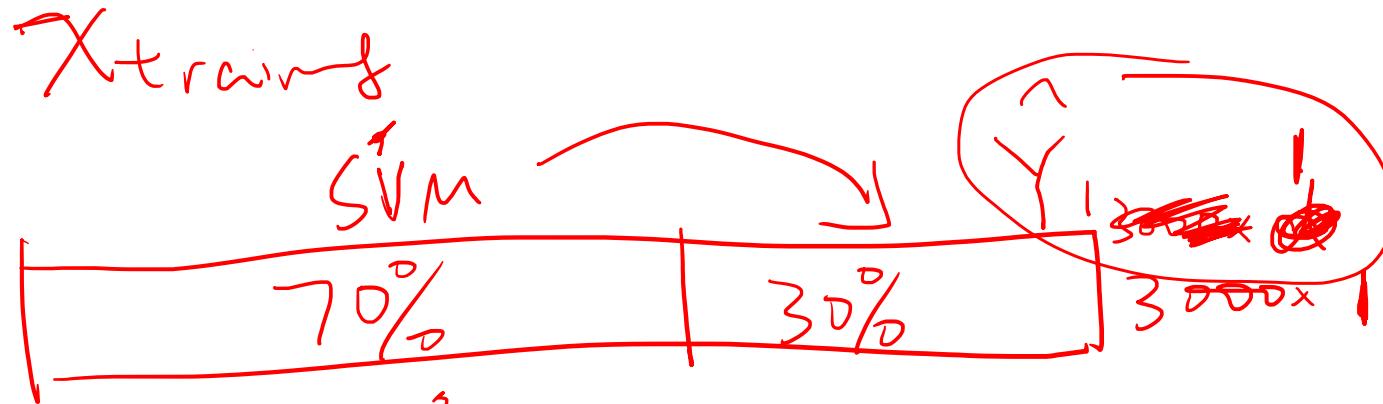






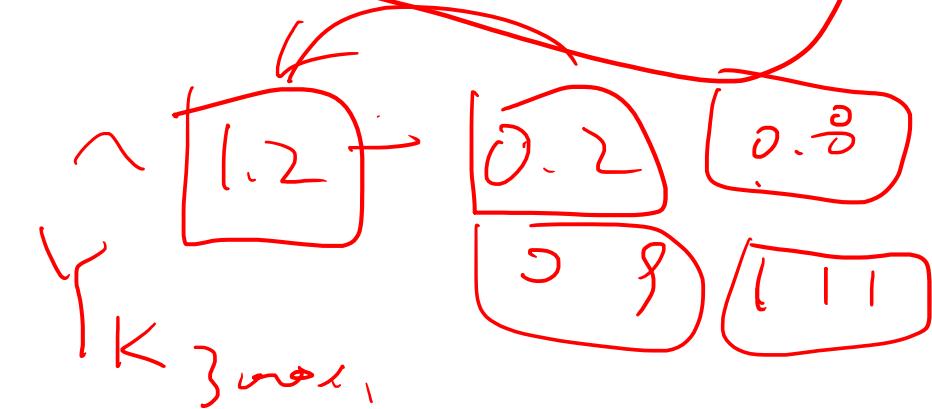






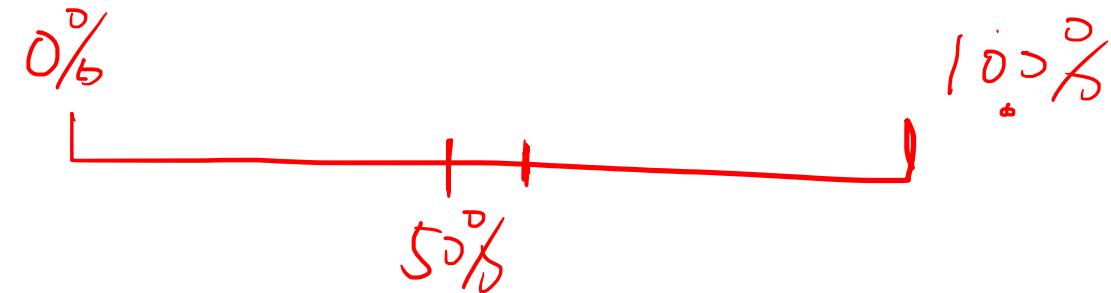
\hat{Y}_3 3000x

$$\hat{Y} = \alpha_1 \hat{Y}_1 + \alpha_2 \hat{Y}_2 + \dots + \alpha_n \hat{Y}_n$$



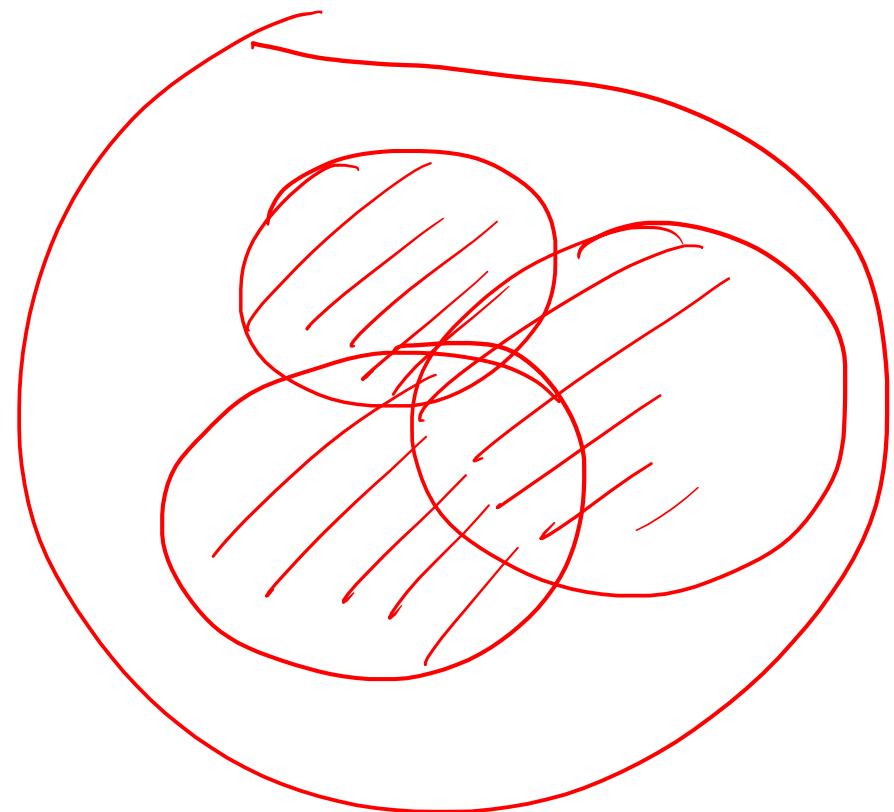
adaboost

$$h_i \in [-1, +1]$$



$$H = \text{Sign} (h_1 + h_2 + \dots + h_n)$$

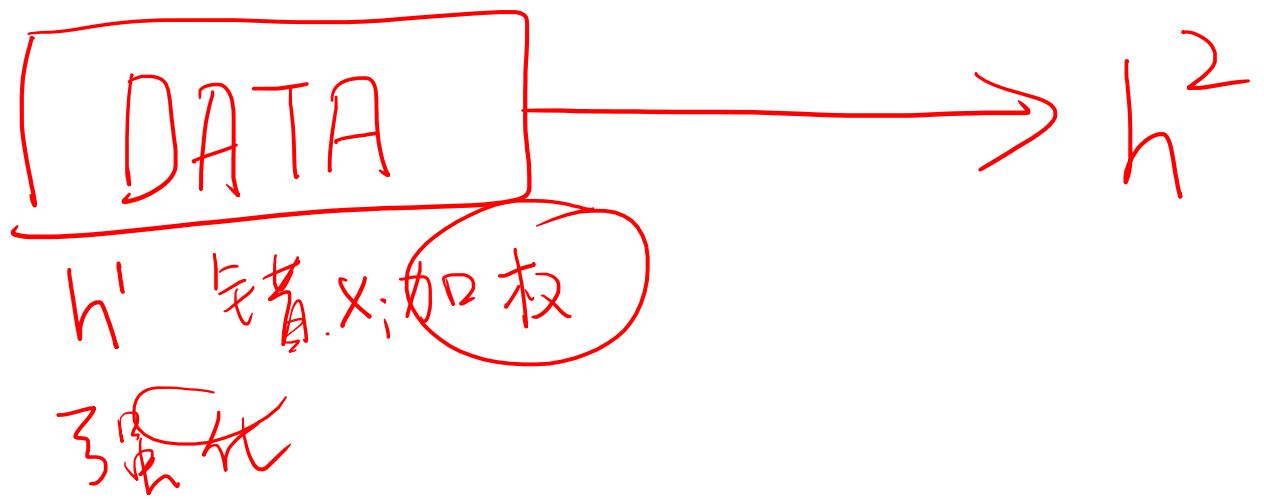
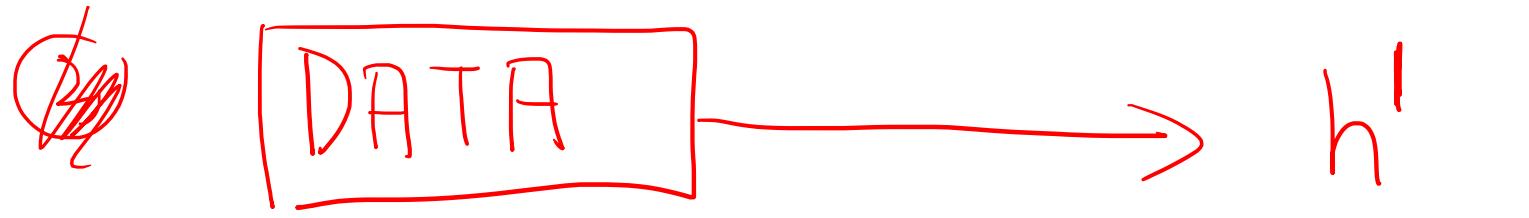
A diagram illustrating the AdaBoost model. A large circle labeled H represents the final hypothesis, which is the sign of the weighted sum of base hypotheses h_1, h_2, \dots, h_n . Below the main equation, a cloud-like shape contains several smaller circles, each labeled h_1, h_2, h_3 , representing the individual base hypotheses. Arrows point from the labels h_1, h_2, h_3 to their respective circles.



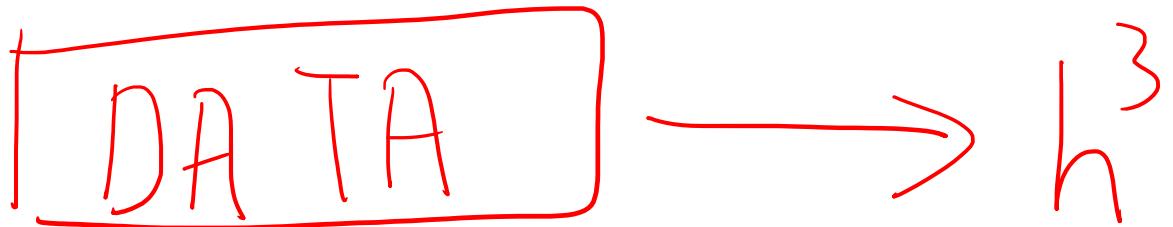
① 対称本加权.

$$\text{LOSS} = \frac{1}{N} \sum_i (\hat{y}_i - y)^2$$

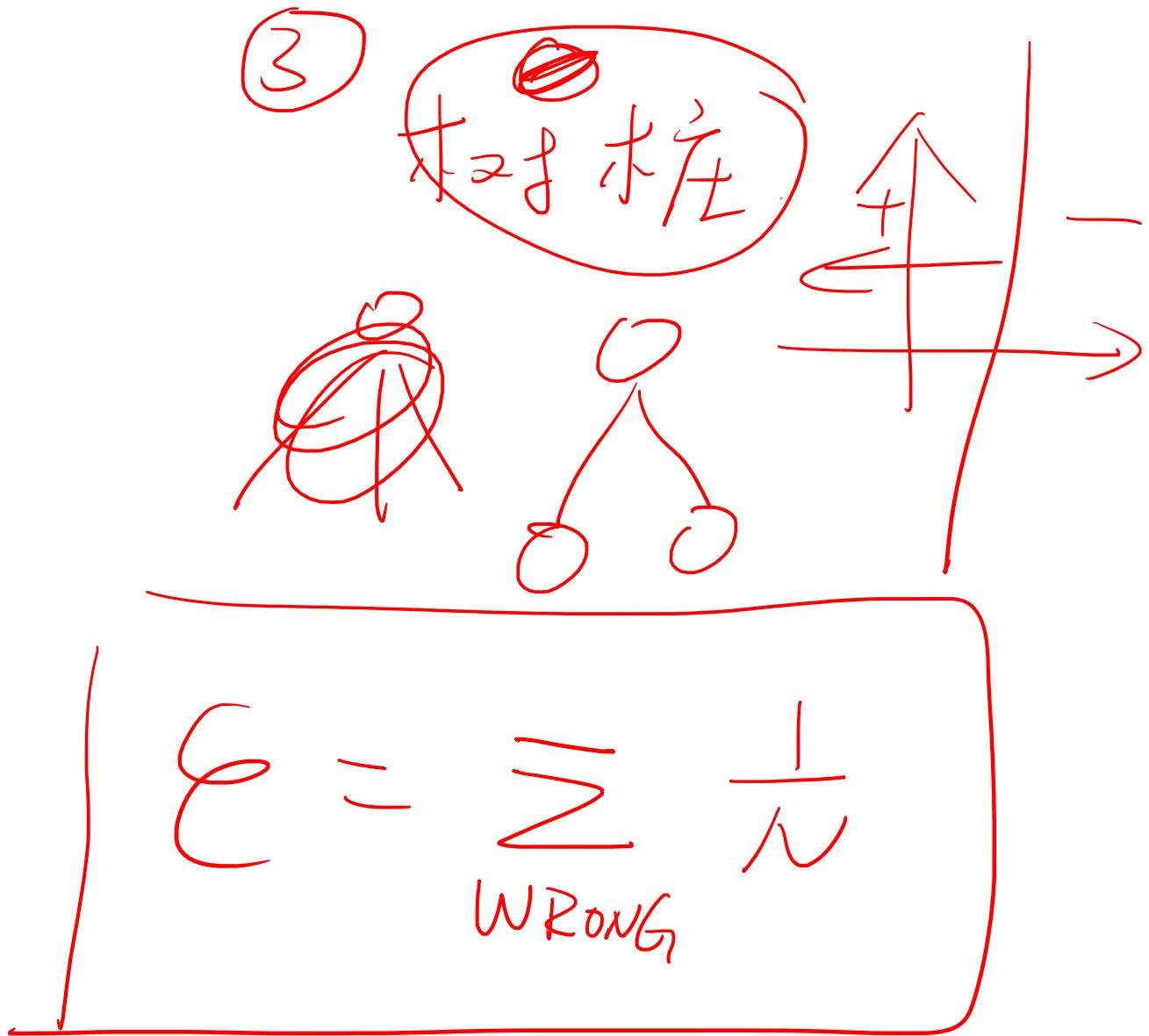
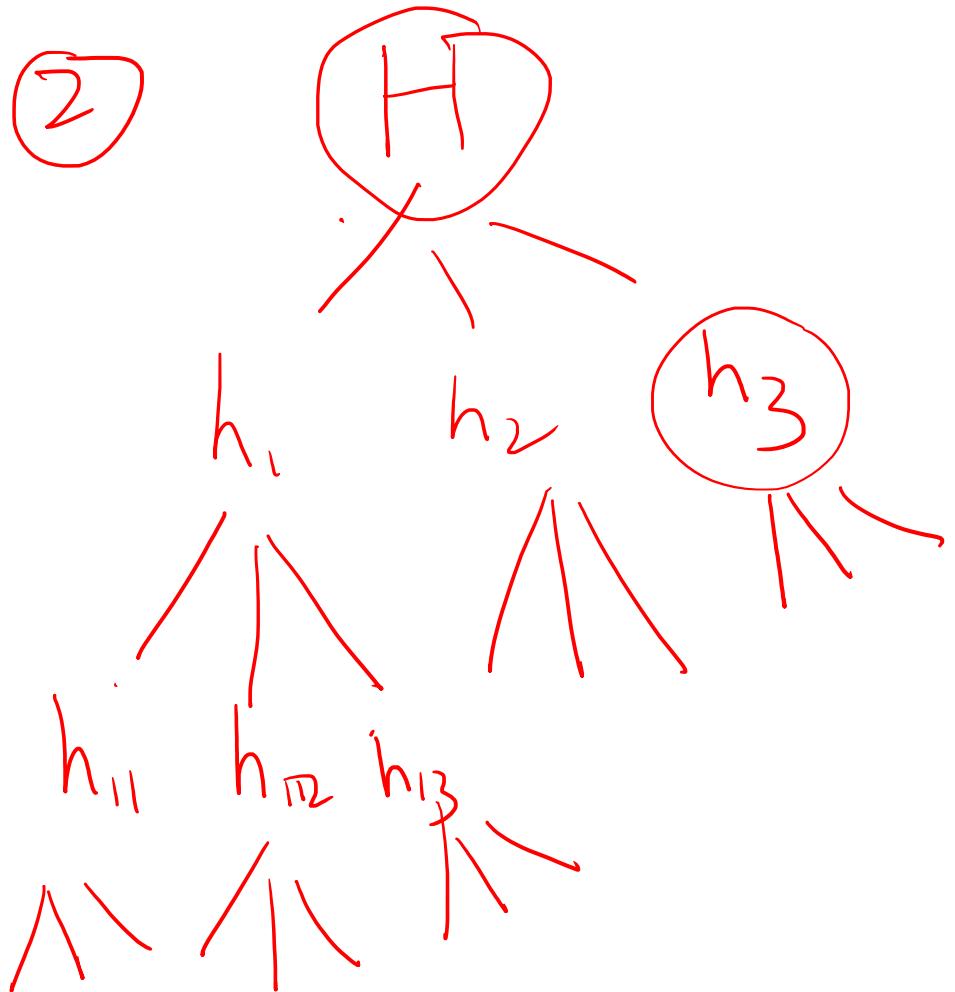
$$\min \text{LOSS}(\theta)$$



$$H = \begin{pmatrix} & h_1 & h_2 & h_3 \end{pmatrix}$$

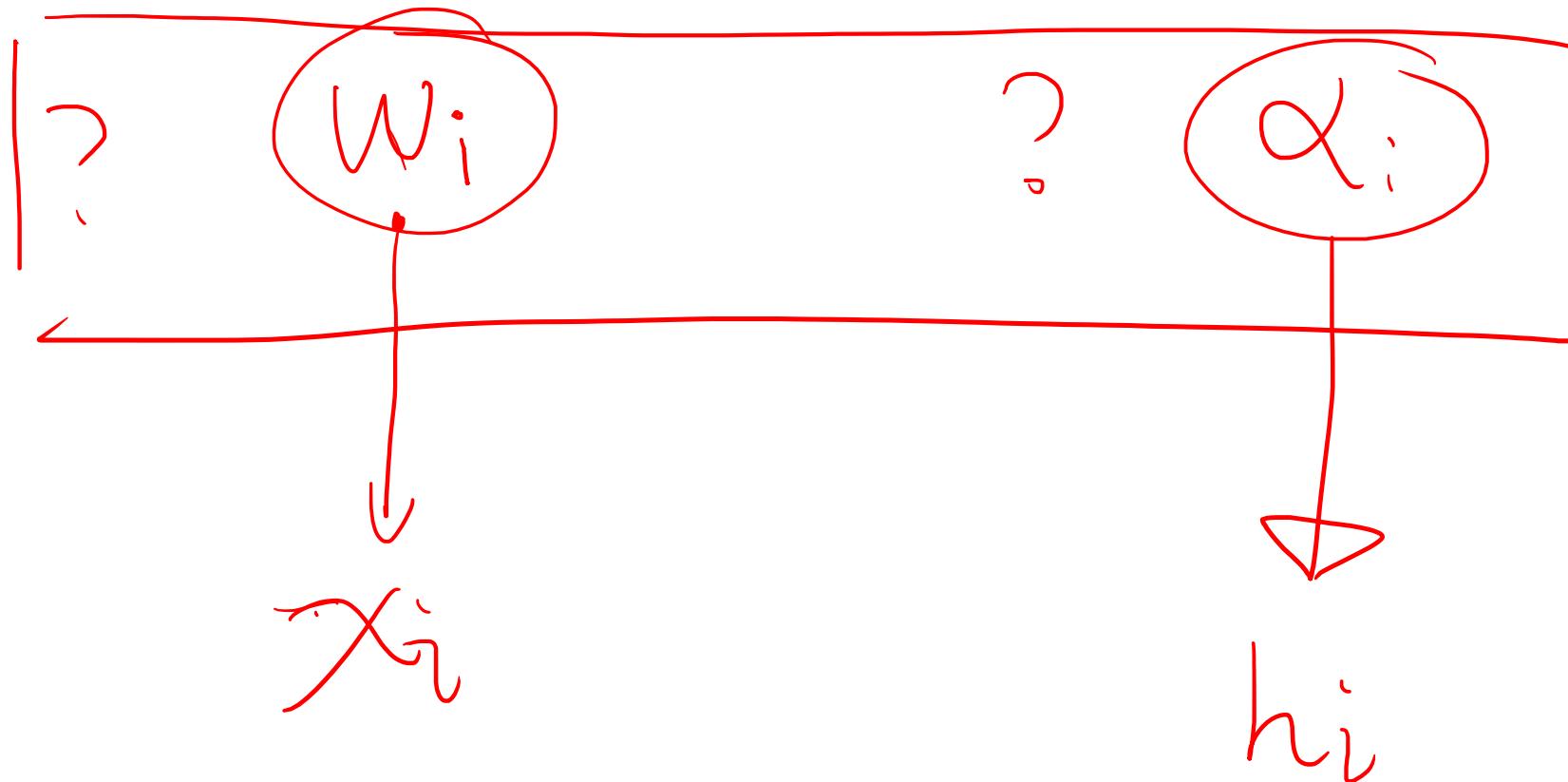


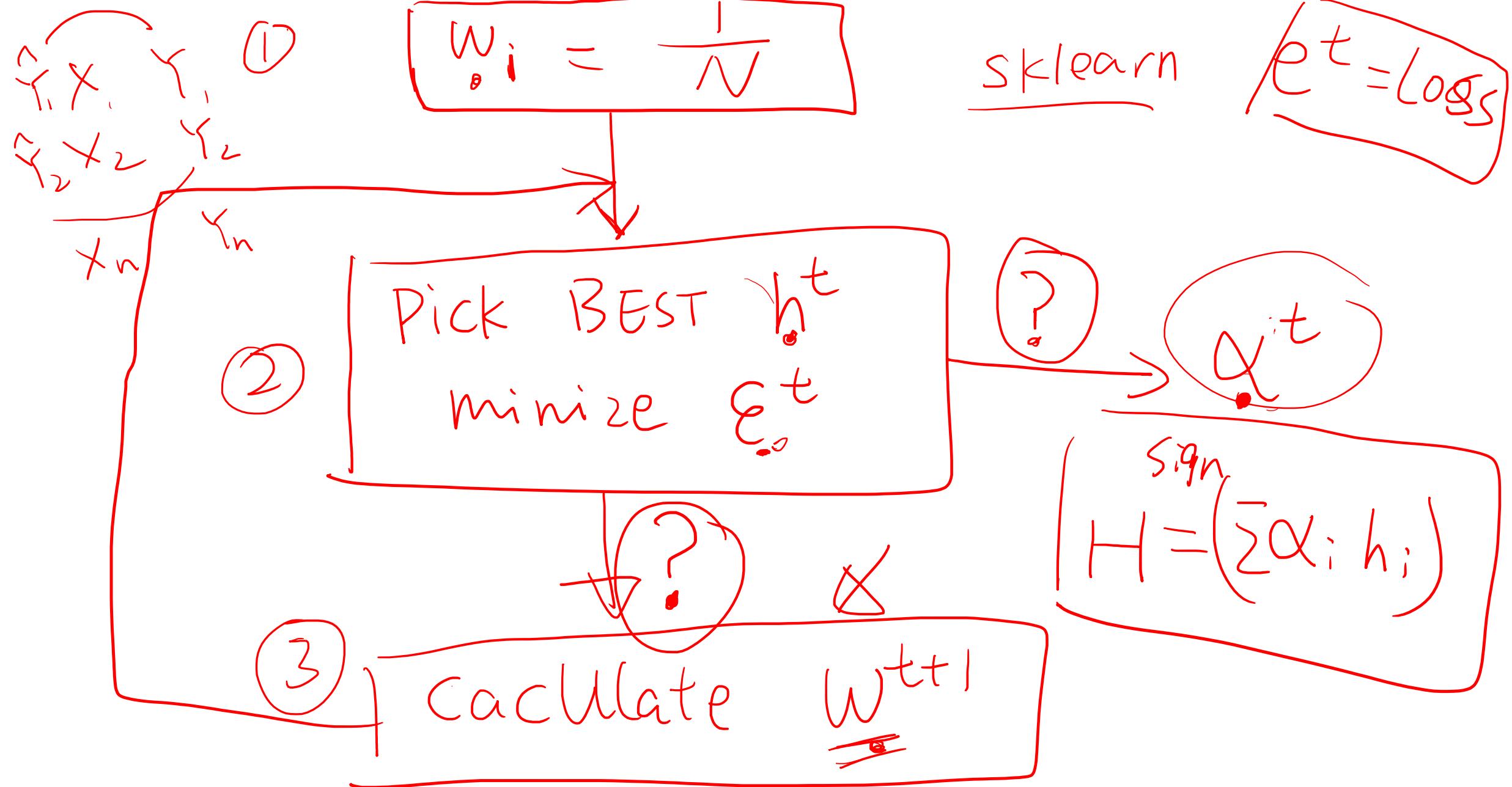
$$h_1 + h_2$$



$$H(x) = \text{Sign}(\alpha_1 h_1 + \alpha_2 h_2 + \dots + \alpha_n h_n)$$

± 1





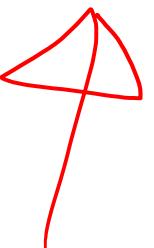
$$w_i^{t+1} = \underbrace{w_i^t \cdot e}_{\text{Normalization}} - \alpha \cdot h(x)^t \cdot y_i$$

± 1 ✓



ERROR BOUND

$$\alpha_t = \frac{1}{\sum} \ln \frac{1 - \epsilon^t}{\epsilon^t}$$



Loss

$$Z = \sqrt{\frac{\epsilon^t}{1-\epsilon^t}} \sum_{\text{correct}} w_i + \sqrt{\frac{1-\epsilon}{\epsilon}} \sum_{\text{WRNG}} -w_i$$

ϵ^t

$$W = \frac{w_i}{z}$$

$$\frac{1-\epsilon}{\epsilon} \times$$

$$Z = 2 \sqrt{\epsilon^t(1-\epsilon^t)}$$

$$W_i^{t+1} = \frac{w^t}{2} + \frac{1}{1-\epsilon}$$

$\sum w_i = \frac{1}{2}$

Correct

$$W_i^{t+1} = \frac{1}{2}$$

WRONG

$$W_i^{t+1} = \frac{w^t}{2} + \frac{1}{1-\epsilon}$$

x ✓

$x_i: \frac{1}{2} \cdot \frac{1}{20} \geq 0.05$

$xx: \frac{1}{2} \cdot \frac{1}{20} \leq 0.05$

$\frac{1}{2} \cdot \frac{1}{20} = 0.05$

$$W_t = \frac{1}{\sum 1 - \epsilon}$$

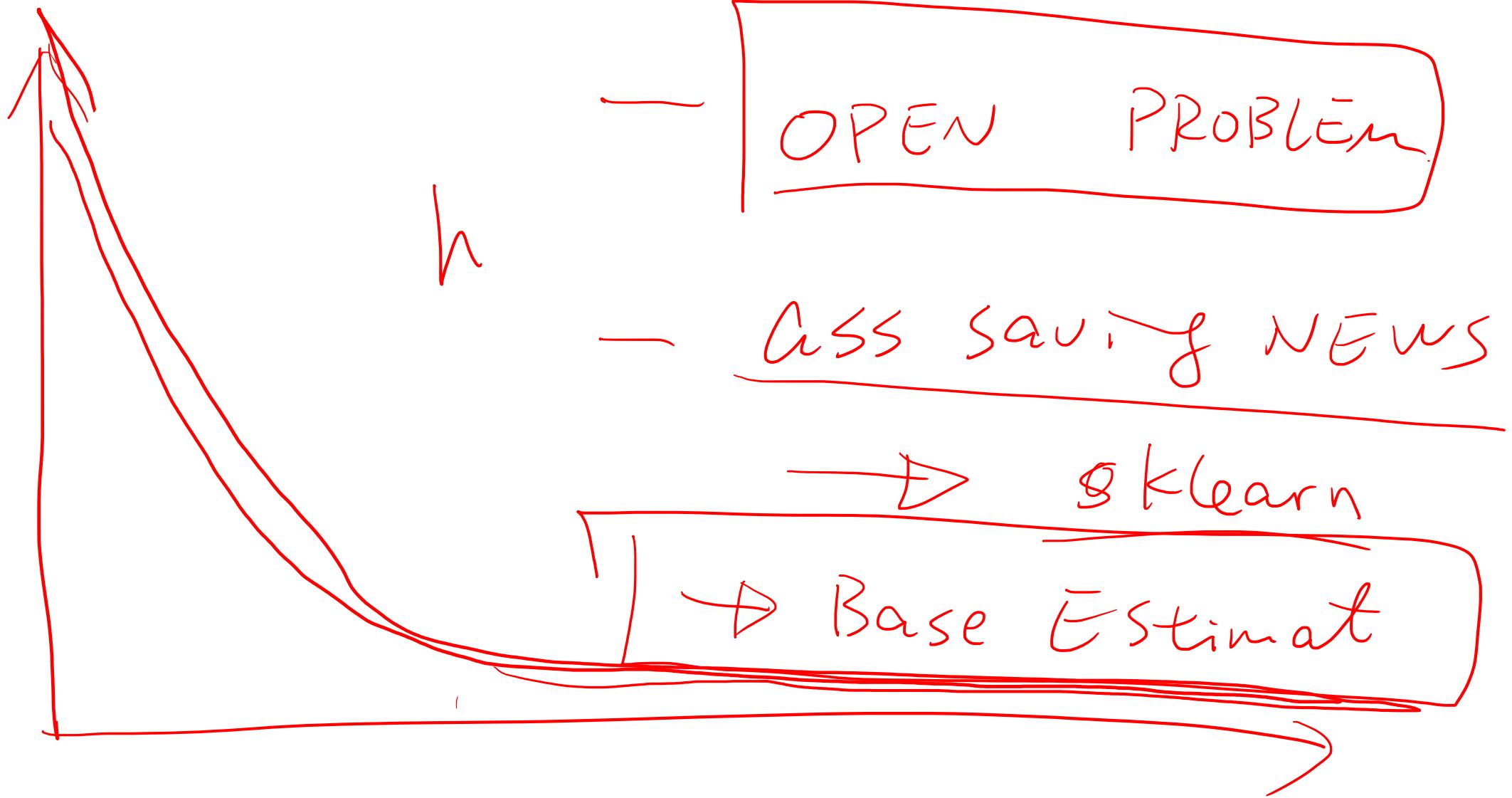
Correct

~~THANK GOD STONE~~

$$\sum w_t$$

Correct

$$= 1 - e^t$$



SKLEARN

